NPS55TY71041A

United States Naval Postgraduate School

SENSITIVITY TESTING

FOR

SAFETY AND RELIABILITY

by

J. B. Tysver

April 1971

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral A. S. Goodfellow, USN
Superintendent

M. U. Clauser
Provost

ABSTRACT:

Probit analysis and the 'staircase' technique for sensitivity testing require prior knowledge or assumptions to provide the location and/or spacing of stimulus levels to be used in the experimentation. A preliminary experimentation phase using a modification of binary search is proposed to provide the input information necessary for efficient use of these techniques. For applications where the experimentation is severely limited, the results of this modified binary search can be used for the purpose of comparison of the reliability and safety of competitive systems.

Prepared by:

J. Bryce Tysver
J. Bryce Tysver
Department of Operations Research
and Administrative Sciences

Approved by:

J. R. Borsting
J. R. Borsting, Chairman
Department of Operations Research
and Administrative Sciences

Released by:

C. E. Menneken
C. E. Menneken
Dean of Research Administration

NPS55TY71041A

I. Introduction.

The term 'sensitivity testing' applies to Bernoulli experiments in which the outcome of each experiment is a quantal response ('0' or '1') to an applied stimulus where the probability, $p(x)$, of a '1' response varies with the stimulus level x. Design and analysis of sensitivity tests are aimed at determining properties of the response function, $p(x)$.

The Probit technique for sensitivity testing has been under development for about a century (Finney [7]). This technique was initially established for applications in the area of bio-assay where, for example, simultaneous treatment of batches of insects by different levels of a toxin can be performed. It is particularly useful when the experimental program can be designed for simultaneous multiple trials at previously specified stimulus levels.

Staircase (Up-and-Down) techniques were introduced in the last twenty five years (Dixon and Mood [5]). These involve sequential experimentation which is usually more practical than batch testing for applications such as failure of physical structures or electronic components subjected to stress, ignition of combustibles subjected to heat, and detonation of munitions or explosives subjected to shock.

The more recent technique of Stochastic Approximation received its initial impetus from the search for efficient procedures for estimating the stimulus level x such that $p(x) = a$ (Robbins and Monro [12]).

Of primary concern in much of the literature is the development of techniques for estimating the stimulus level $x_{.5}$ at which $p(x) = 0.5$ (Lethal Dosage 50 or LD50 in bio-assay terminology). Estimation of standard deviation was included primarily for the purpose of making confidence statements about $x_{.5}$. More recent papers recognize the standard deviation of the stimulus level as a primary sensitivity measure to be distinguished from its secondary role in establishing confidence intervals for $x_{.5}$ (Gayle [8]). The discussion in Section III illustrates the fact that knowledge of $x_{.5}$ is not, in general, sufficient information for establishing safety and reliability of a system nor for comparison of those factors for different systems.

Application of the Probit and Staircase techniques require prior knowledge or assumptions by the investigator for the specification of the location and/or spacing of the applied stimulus levels. In addition, it is usually assumed that the stimulus, or some transformation of it, obeys the normal distribution law (Bartlett [2] and Garwood [9]). An investigator may not always be willing, or able, to provide these inputs without some preliminary experimentation. The modified binary search procedure proposed in Section IV is designed for this preliminary search phase with the specific intent of providing a basis for specification of the location and spacing of stimulus levels to be used in the main experimentation program. The proposed search procedure is intended as a prelude, to rather than a substitute for, application of the Probit or Staircase techniques. It could, however, be used for the purpose of comparing alternative systems when the experimentation is severely restricted in the number of trials that can be made. The possibility of using this technique as a basis for an experimentation procedure for the subsequent major experimentation should not be ignored (Evans, et al [6]).

The assumption of normality has been the subject of considerable discussion in the literature (Anscombe [1]). Extensive experimentation is required to establish the appropriate distribution for applications in which safety and/or reliability are of primary interest since these factors involve the tails of the distribution. In the vicinity of $x_{.5}$, the validity of the assumption of normality does not appear to be critical, particularly when the experimentation is limited to small samples ([5] p. 112 and Tysver [13]). In the analysis accompanying the search procedure proposed in Section IV and its extension in Section V, the sample sizes used for estimation are between 6 and 9 and hence the assumption of normality is considered to be acceptable for the stimulus level itself or that an appropriate transformation (such as the logarithm) has been introduced and the stimulus levels are measured in this transformed variable.

The model used in the development and discussion reported in this paper is presented in Section II.

II. The Model.

Let x denote an applied stimulus level $(0 \leq x < \infty)$ and $y = y(x)$ the corresponding quantal response ('0' for no response and '1' for a positive response). At any specified stimulus level the response will be considered to be a realization of a Bernoulli random variable Y with the response probability

$$p(x) = \text{Prob } (Y = 1 \mid x).$$

The function $p(x)$ is called the response function. In general we will have $p(0) = 0$ and $p(\infty) = 1$. A shorter interval $(a \leq x \leq b)$ with $0 \leq a < b < \infty$ can frequently be considered with $p(a) \simeq 0$ and $p(b) \simeq 1$. A response function is sketched in Figure 1.

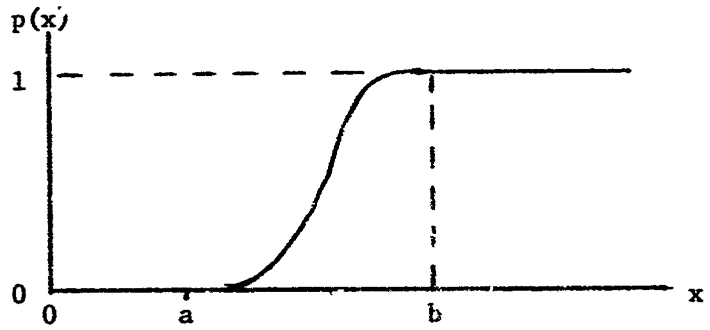


Figure 1. The Quantal Response Curve

Note that

$$\text{Prob } (Y = 1 \mid x \le a) \simeq 0 \quad \text{and} \quad \text{Prob } (Y = 1 \mid x \ge b) \simeq 1.$$

For sensitivity testing applications the response function will be continuous and monotonely increasing in x. Thus $p(x)$ can be considered as the cumulative distribution function for a random variable X with

$$p(x) = \text{Prob } (X \le x).$$

The random variable X will be interpreted as a threshold stimulus level. If the applied stimulus level x satisfies the inequality $x \ge X$ where X is the applicable threshold stimulus level in a particular experiment, then the response will be $y = 1$. If $x < X$, then the response will be $y = 0$. Thus

$$\text{Prob } (Y = 1 \mid x) = \text{Prob } (X \le x) = p(x)$$

and

$$\text{Prob } (Y = 0 \mid x) = \text{Prob } (X > x) = 1 - p(x).$$

In conjunction with the distribution function $p(x)$ there will be a probability density function $f(x)$ for the threshold stimulus X such that

$$p(x) = \int_0^x f(u)du.$$

Sensitivity testing differs from ordinary random sampling in that the experimentation provides data on the distribution function instead of the density function. More specifically, each trial does not provide an observation on the random variable X. It only establishes whether the value for that random variable is greater or less than the applied stimulus level in that particular trial. Less information is obtained from each trial than would be achieved by an observation on the random variable X (the threshold stimulus level). The mathematical form of the density function and estimates for its parameters must be established from this data on the distribution function.

For the analysis of the search procedure described in Sections IV and V, it will be assumed that X obeys the normal distribution. Then

$$p(x) = N(X \mid \mu, \sigma^2)$$

where $N(X \mid \mu, \sigma^2)$ denotes the cumulative normal distribution with mean μ and variance σ^2. In Particular

$$\text{Prob } (X \leq \mu) = p(\mu) = N(\mu \mid \mu, \sigma^2) = 0.5.$$

Two additional concepts will be used in the next section. One is that of a 'mixed response region,' (Golub and Grubb [1?]). If a stimulus level X satisfies the inequality

$$\varepsilon < p(x) < 1 - \varepsilon$$

for reasonably small ε, then x is said to be in the region of mixed responses. One of the goals of a preliminary search procedure is the location of the mixed response region. The other concept deals with the experimental evidence indicating when stimulus levels are in the mixed response region. An 'inversion' will be said to have occurred when two applied stimulus levels x_1 and x_2 with $x_1 < x_2$ produce responses $y(x_1) = 1$ and $y(x_2) = 0$. When this inversion occurs, both x_1 and x_2 are in the mixed response region.

III. Safety and Reliability.

Sensitivity applications can be divided into two types for safety and reliability considerations. The first type includes detonation and ignition applications. The second type includes the bio-assay applications. These are described below.

The first type involves the response of a subject to two different stimulus sources. Let $f_t(x)$ denote the density function of the threshold response stimulus for a subject (e.g., a sample of explosive material). Let $f_s(x)$ denote the density function for the signal stimulus applied when a response is desired. Reliability considerations involve the relationship between $f_t(x)$ and $f_s(x)$. Similarly, let $f_n(x)$ denote the density function for the background or noise stimulus level. Safety considerations involve the relationship of $f_t(x)$ to $f_n(x)$. These relationships are illustrated in Figure 2.

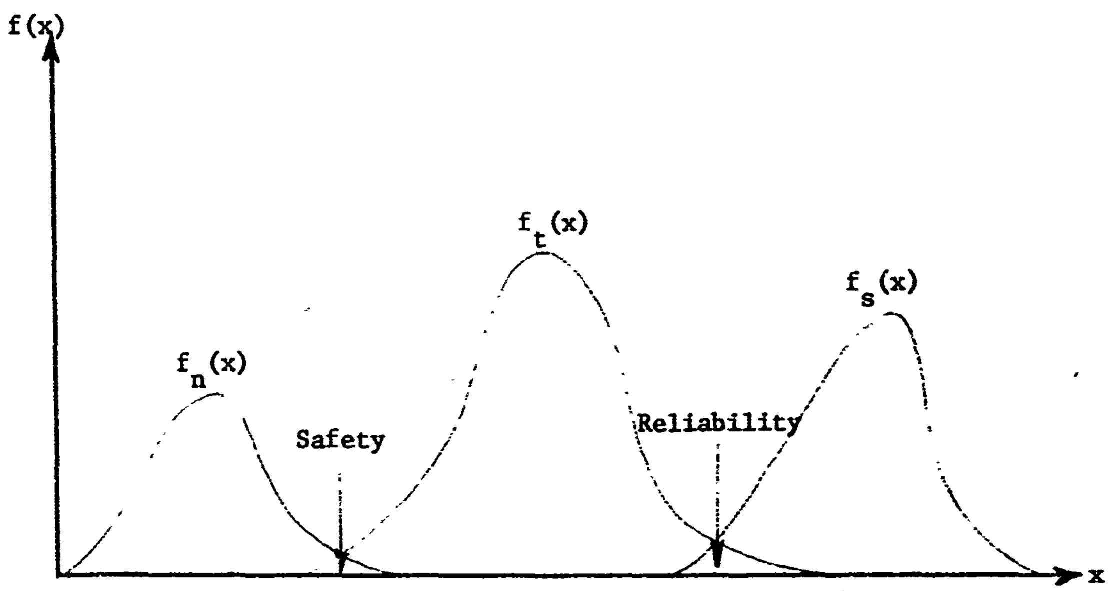


Figure 2. Relationship of Response Threshold to Signal and Noise Stimuli for Detonation and Ignition Applications.

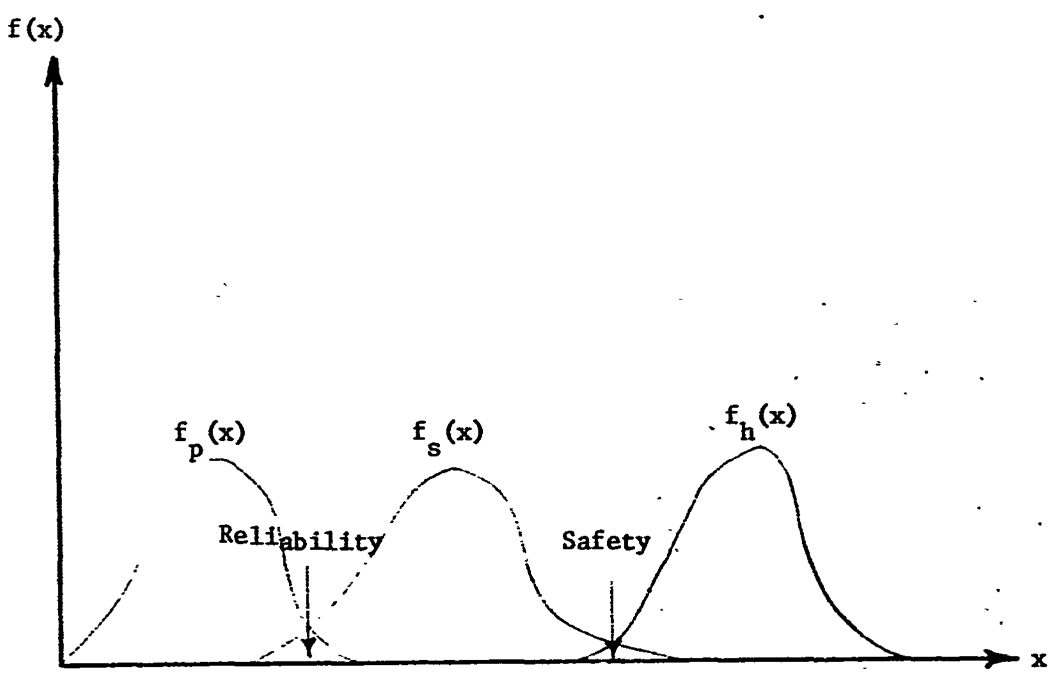


Figure 3. Relationship of Signal Stimulus to Host and Parasite Response Thresholds for Bio-Assay Applications.

A safety measure for a particular application can be established as follows. Let

$$F_t(x) = \int_0^x f_t(x)du.$$

The probability that an undesired response will occur is the probability that a response will be caused by noise, i.e.,

$$P(Y = 1/\text{noise}) = \int_0^\infty F_t(x)f_n(x)dx.$$

The probability that no undesired response will occur can be used as a safety measure.

$$S = 1 - P(Y = 1/\text{noise}).$$

This represents an extreme simplification of safety considerations since the noise stimulus level can vary with time and such factors as the method and, for example, the extent of transportation of an explosive material prior to the time a response is desired.

In a similar way, a reliability factor can be defined as the probability that a response will occur when a Signal stimulus is applied, i.e.,

$$R = P(Y = 1/\text{signal})$$
$$= \int_0^\infty F_t(x)f_s(x)dx.$$

The second type of application involves the response of two subjects to an applied stimulus. When a signal stimulus is applied, a response is desired from one of the subjects (called the parasite) but not from the other (the host).

Let

$f_s(x)$ = density function for applied stimulus,

$f_p(x)$ = density function for threshold response level of parasite,

$f_h(x)$ = density function for threshold response level of host.

It should be noted that, since high safety and reliability measures are usually desired, one or both tails of density functions for threshold response stimuli are involved. The stimulus level $x_{.5}$ at which the probability of a response by a subject is of interest for safety or reliability considerations only in its relationship to the tails of the density function for the related threshold response stimulus.

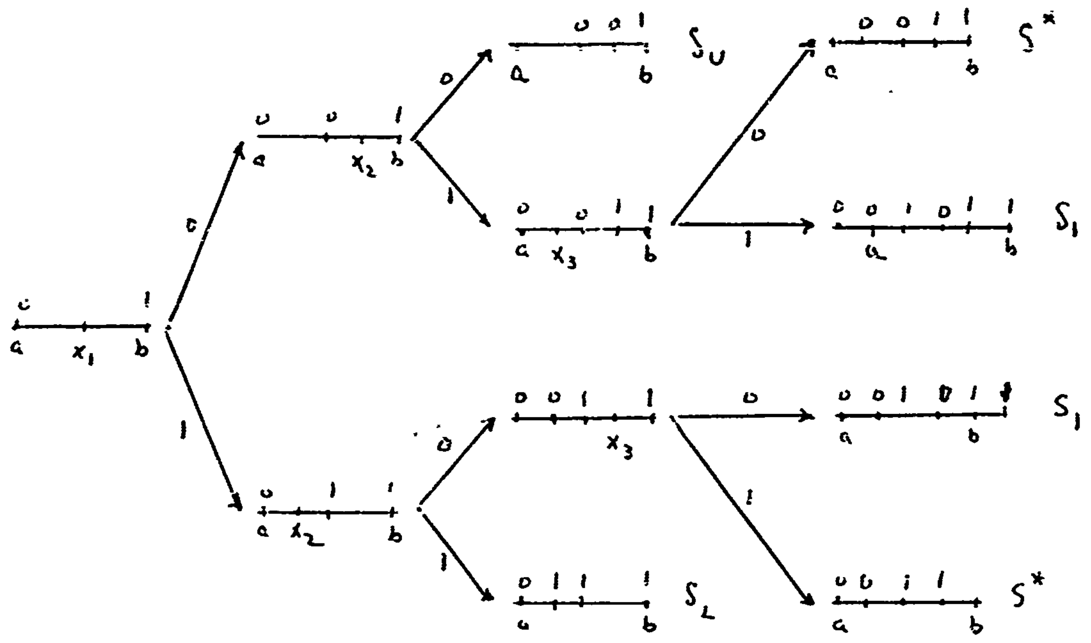
To illustrate the relationship of $x_{.5}$ to safety and reliability consider an application to explosives. If the physical or chemical properties of an explosive are changed to shift the threshold density function $f_t(x)$ to the right without changing its shape, (increasing μ_t) the safety measure will be increased. The applied signal density function $f_s(x)$, which is usually subject to control by the operator or engineer in charge, must then be shifted to the right also (increasing μ_s) to prevent decrease in the reliability measure.

On the other hand, if changes can be made to reduce the spread of $f_t(x)$ (reducing σ_t) both safety and reliability measures could be improved. An investigator should be interested in establishing the effect of changes on σ_t as well as on μ_t. It is a conjecture of the author that changes in the composition of an explosive to increase μ will also increase σ_t. This increase in σ_t could at least partially counteract the gain in safety produced by the increase in μ_t and would necessitate an additional increase in the mean μ_s of the applied signal to maintain the reliability.

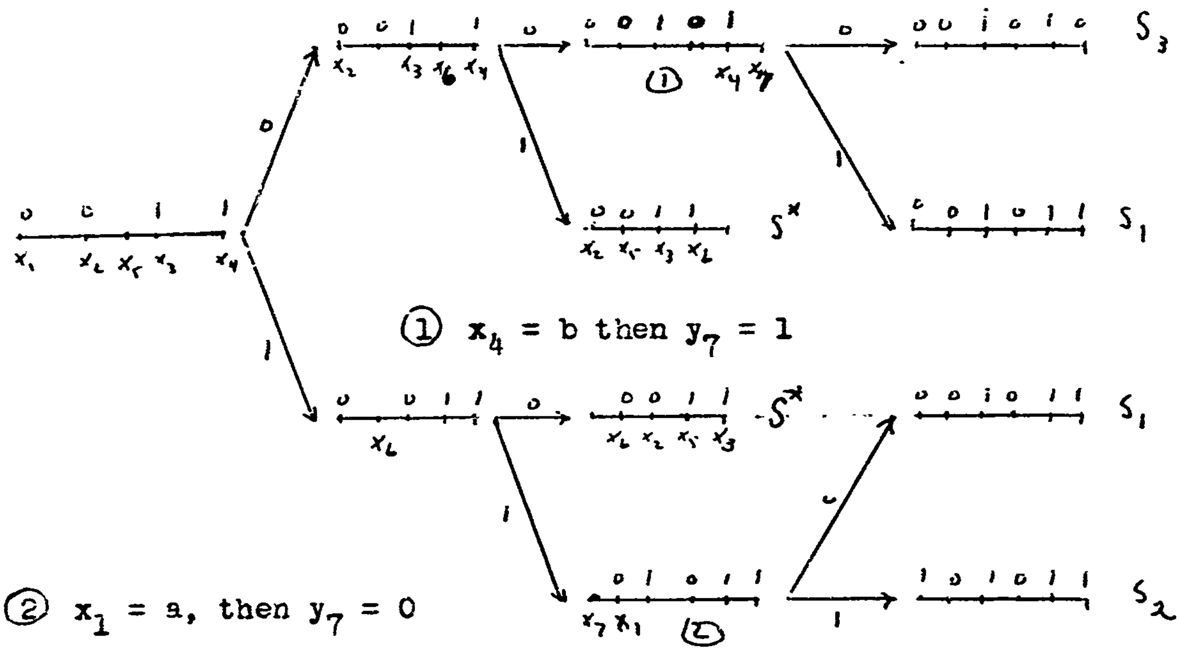
IV. A Modified Binary Search.

If the response curve represents a step function, then a simple binary search for the location of the step would be appropriate. When the mixed response region has non-zero width, however, the occurrence of a '0' ('1') response does not indicate whether the applied stimulus is below (above) or in that region. The binary search procedure must then be modified to provide some indication when applied stimulus levels are in the mixed response region. Response inversion is the indicator incorporated in the proposed procedure shown in the flow charts (Figure 4) and described below.

Situation S* is a cyclic one indicating that a reduction in step size should be introduced. The procedure aims at reconstruction of this situation at each cycle. Failure to reconstruct S* occurs when there is a response inversion indicating that the

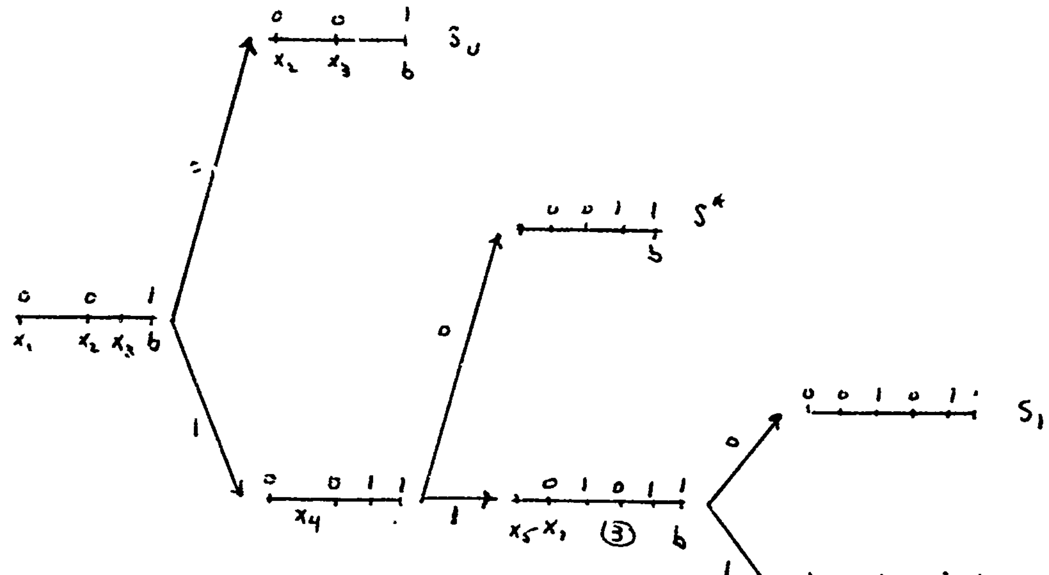


a. Start of Search



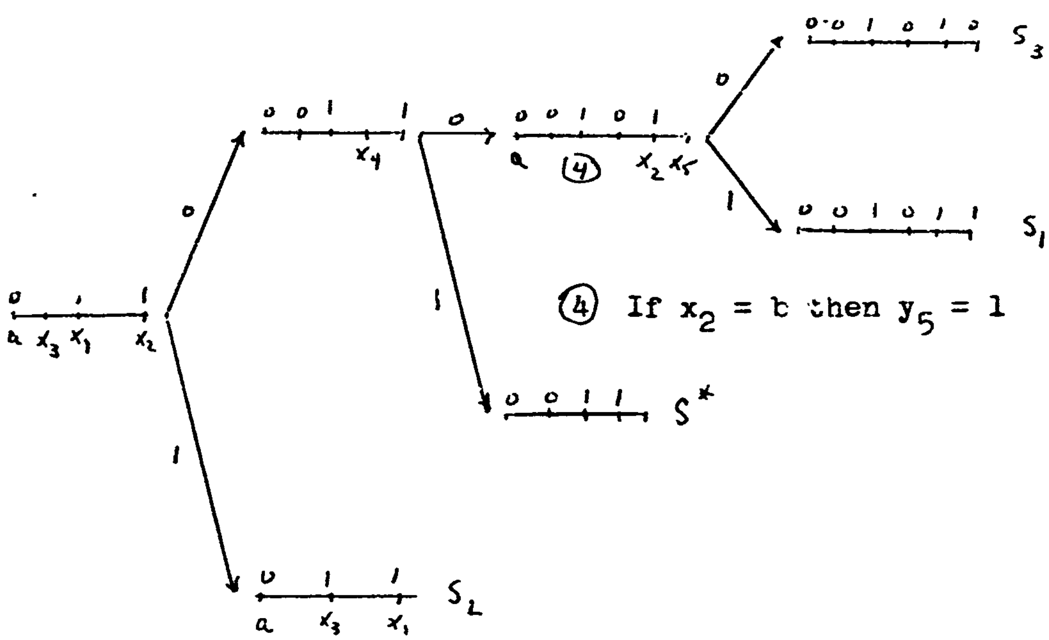
b. Search from S*

Figure 4. Flow Chart for Modified Binary Search



③ If $x_1 = a$ then $y_5 = 0$

c. Search near Upper Boundary (S_U)



④ If $x_2 = b$ then $y_5 = 1$

d. Search near Lower Boundary (S_L)

Figure 4. (Continued)

general location of the mixed response region has been established. Trials are then introduced at the ends of the basic inversion situation S_0 to produce the terminal situations S_1, S_2 or S_3 (Figure 5). S_U and S_L are upper and lower boundary situations, respectively. Footnotes in Figure 4 indicate that the proposed sequence of trials could indicate trials to be made at a or b. The confidence in the assumed responses $(y(a) = 0$ and $y(b) = 1)$ by the investigator should indicate whether these trials are necessary.

In the analysis which follows, a normal distribution will be assumed for the threshold stimulus and trials at different stimulus levels will be assumed to be independent. Thus, for example, the probability of the responses in situation S^* is

$$\text{Prob }(S^*) = \text{Prob }(Y_1 = 0, Y_2 = 0, Y_3 = 1, Y_4 = 1 | x_1, x_2, x_3, x_4)$$

$$= \prod_{i=1}^{4} \text{Prob }(Y_i = y_i | x_i) = [1 - N(x_1)][1 - N(x_2)]N(x_3)N(x_4)$$

where

$$\text{Prob }(Y_i = y_i | x_i) = \begin{cases} N(x_i) & \text{if } y_i = 1 \\ 1 - N(x_i) & \text{if } y_i = 0 \end{cases}$$

and

$$N(x_i) = \text{Prob }(X_i \leq x_i) = \int_{-\infty}^{x_i} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

Figure 5 includes the situations of primary interest.

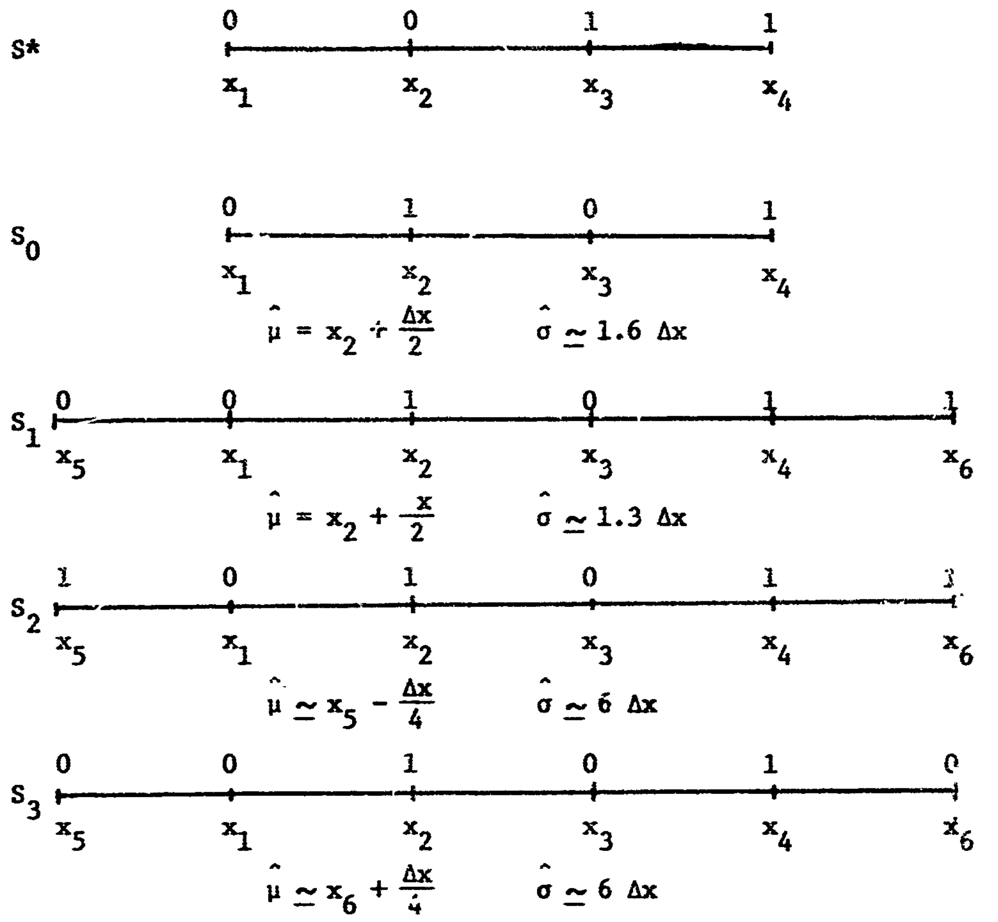


Figure 5.

Experimental Situations

Maximum likelihood estimates for μ and σ can then be established using tables for the normal distribution for each of the terminal situations (see Golub and Grubb [10] and Langlie [11]). Only approximate values for the estimates $\hat{\mu}$ and $\hat{\sigma}$ are required from the preliminary search phase.

It may be of some interest to consider a general situation S consisting of n trials at equally spaced stimulus levels. If μ is within the sample and σ is large with respect to the sample width then the responses at the individual stimulus levels are immaterial since

$$\text{Prob } (Y = 0|x) \simeq \text{Prob } (Y = 1|x) \simeq 0.5.$$

Then

$$\text{Prob } (S) \simeq (0.5)^n,$$

$$\text{Prob } (S^*) \simeq \text{Prob } (S_0) \simeq (0.5)^4 = 0.0625,$$

and

$$\text{Prob } (S_1) \simeq \text{Prob } (S_2) \simeq \text{Prob } (S_3) \simeq (0.5)^6 \approx 0.0156.$$

The analysis for S^* is simple. By symmetry, $\hat{\mu} = x_2 + \Delta x/2$ where $\Delta x = x_i - x_{i-1}$. Also Prob (S^*) approaches unity as σ approaches zero. Thus a reduction in Δx is indicated with the region of mixed responses lying between x_2 and x_3. For S_0, symmetry also holds so that $\hat{\mu} = x_2 + \Delta x/2$. The estimate $\hat{\sigma} \simeq 1.6\Delta x$ together with $\hat{\mu}$ gives Prob $(S_0 \simeq 0.097$, which is not substantiallly greater than the limiting value Prob $(S_0|\sigma >> \Delta x) \simeq 0.0625$. Prob (S_0) does not change appreciably with substantial changes in μ and σ as shown in Table 1. (This could be anticipated since only 4 trials are involved.) Since the estimate of σ is greater than Δx, additional trials should be used in an attempt to pin down the ends of the mixed response region instead of reducing the step size further.

Situation S_1 also involves symmetry so that $\hat{\mu} = x_2 + \Delta x/2$. Further, $\hat{\sigma} = 1.3\ \Delta x$ and Prob $(S_1|\hat{\mu},\hat{\sigma}) \simeq 0.089$, which is substantially larger than the limiting value Prob $(S_0|\sigma >> \Delta x) = 0.0156$ (see Table 2). S_1 is accepted as a terminal situation, but this should not preclude an investigator from taking additional trials to improve the estimates. Situations S_2 and S_3 are similar so only one need be analyzed. The estimates for S_2 are

$$\hat{\mu} \simeq x_5 - \frac{\Delta x}{4} \quad \text{and} \quad \hat{\sigma} \simeq 6\ \Delta x$$

For S_3 they are

$$\hat{\mu} \simeq x_6 + \frac{\Delta x}{4} \quad \text{and} \quad \hat{\sigma} \simeq 6\ \Delta x.$$

If an investigator wishes, previous responses to stimulus levels which are near the terminal situations can be included in establishing the maximum likelihood estimates. Trials substantially

TABLE 1

$P(S_0|\mu,\sigma)$

S_0	0	1	0	1
	x_1	x_2	x_3	x_4

σ \ μ	x_2	$x_2 + \frac{\Delta x}{4}$	$x_2 + \frac{\Delta x}{2}$	$x_2 + \frac{3\ \Delta x}{4}$	x_3
Δx	.06667		.08288		.06667
$1.5\ \Delta x$	.08577	.09376	.09590	.09376	.08577
$1.6\ \Delta x$	.08729	.09428	<u>.09728</u>	.09428	.08729
$1.75\ \Delta x$	.08881	.09491	.09709	.09491	.08881
$2\ \Delta x$	.08974	.09463	.09633	.09463	.08974
$2.5\ \Delta x$	.08900	.09217	.09321	.09217	.08900

By symmetry $\quad \hat{\mu} = x_2 + \frac{\Delta x}{2}$

$(\sigma >> \Delta x, \mu = \hat{\mu}) \Rightarrow P(S_0) \simeq (.5)^4 = .0625$

$\hat{\sigma} = \max_{\sigma} P(S_0|\hat{\mu},\sigma) \simeq 1.6\ \Delta x$

TABLE 2

$P(S_1|\mu,\sigma)$

σ	$\frac{\Delta x}{2}$	Δx	$1.25\Delta x$	$1.3\Delta x$	$1.4\Delta x$	$1.5\Delta x$	$1.6\Delta x$	$1.75\Delta x$	$2\Delta x$	$2.5\Delta x$	$>>\Delta x$
$P(S_1)$	.025	.082	.0888	.0890	.0887	.087	.086	.083	.077	.066	.016

$\mu = \hat{\mu} = x_2 + \frac{\Delta x}{2}$ $\hat{\sigma} \simeq 1.3\ \Delta x$

TABLE 3

$P(S_2|\mu,\sigma)$

S_2	1	0	1	0	1	1
	x_5	x_1	x_2	x_3	x_4	x_6

μ \ σ	$4\ \Delta x$	$5\ \Delta x$	$6\ \Delta x$	$7\ \Delta x$
x_2	.01982	.01997		
$x_1 + \frac{\Delta x}{2}$	.02252	.02230		
x_1	.02411	.02402	.02345	
$x_1 - \frac{\Delta x}{2}$	.02452	.02497	.02459	
$x_5 + \frac{\Delta x}{4}$	.02424	.02516		
x_5		.02507	.02516	.02470
$x_5 - \frac{\Delta x}{4}$			$\underline{.02522}$	.02500
$x_5 - \frac{\Delta x}{2}$		.02434	.02514	.02510
$x_5 - \frac{3\Delta x}{4}$				.02511
$x_5 - \Delta_x$				.02498

$\hat{\mu} \simeq x_5 - \frac{\Delta x}{4}$ $\qquad$ $\hat{\sigma} \simeq 6\ \Delta x$

$(\mu >> \Delta x, \mu = \hat{\mu}) \Rightarrow P(S_2) \simeq (.5)^6 \simeq .0156$

separated from those in the terminal situations will contribute little to the estimates.

V. Extension of Search.

The limitation in Section IV to the responses at the six stimulus levels of S_1, S_2 and S_3 for the estimation of μ and σ appear reasonable for many applications where experimentation must be severely limited. For the purpose of comparison of fuels or explosives, some investigators may consider that the preliminary search procedure provides sufficient evidence without any subsequent experimentation. A substantial portion of the literature on sensitivity testing is devoted to sample sizes of 10 or less. If the investigator can allocate more trials to the preliminary search phase, the proposed search procedure can be easily extended. Some suggested extensions are shown in Figures 6, 7 and 8.

For situation S_1, the ends of the mixed response region appear to be pinned down and the additional stimulus levels should be used to fill in between the levels already used. Two suggestions involving 3 or 5 additional levels are shown. Depending on the investigator's resources, these could be added directly as indicated by the upper and lower paths in the diagram, or sequentially. If the sequential path is chosen, the three central levels are used first and then a decision is made whether or not to add the other two. If the responses to the first three are (0,0,1) or (0,1,1) the other two may be dropped. However, if any of the other response sequences occur, the other two should be used.

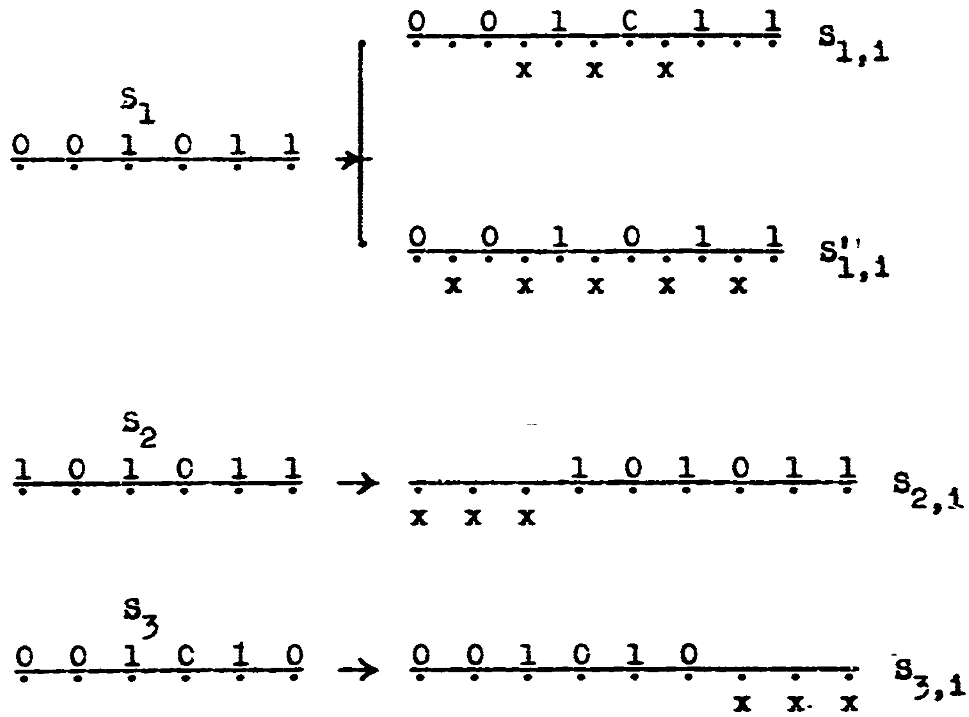


Figure 6. Extension of Search

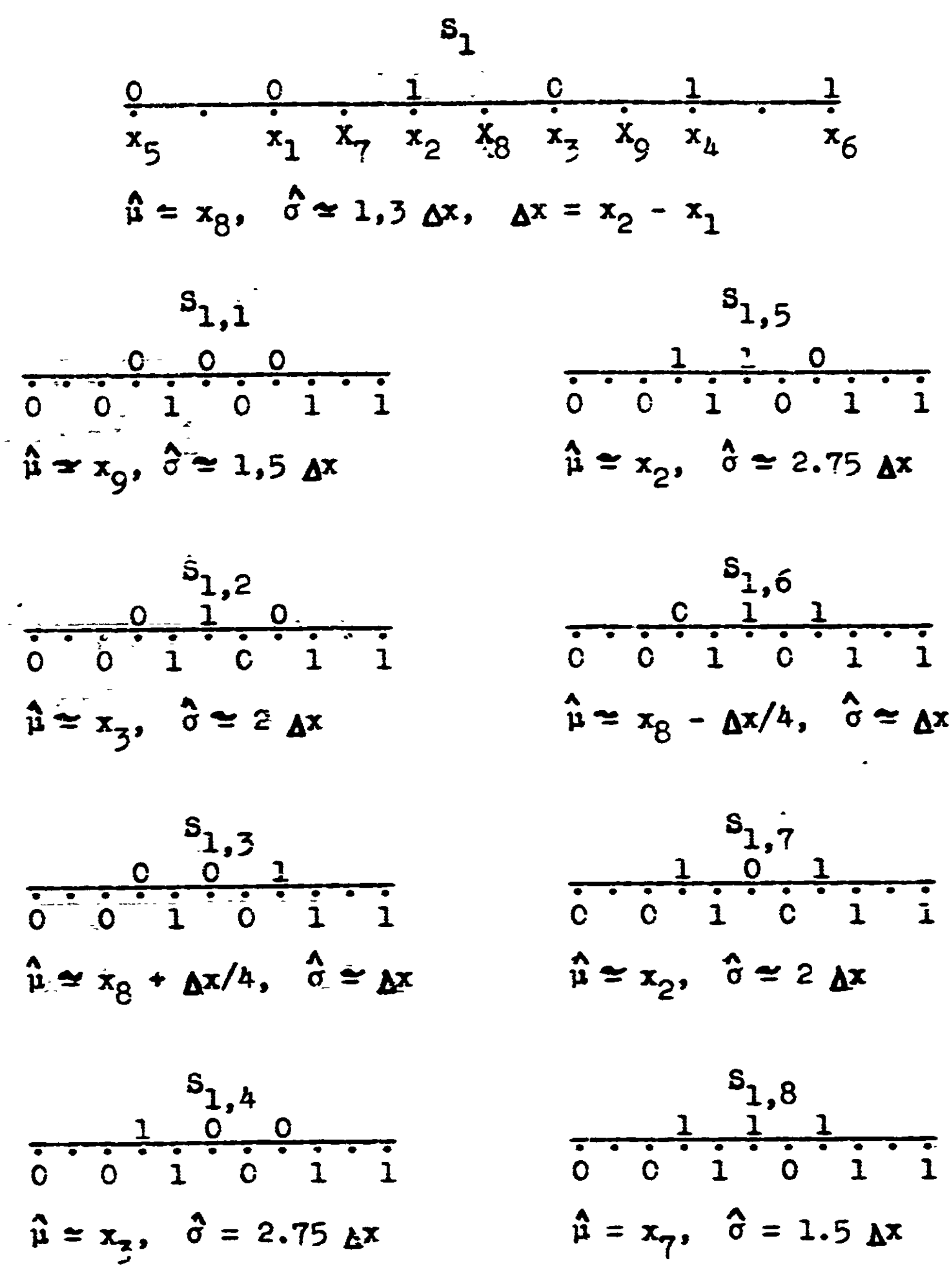


Figure 7. Extension of S_1

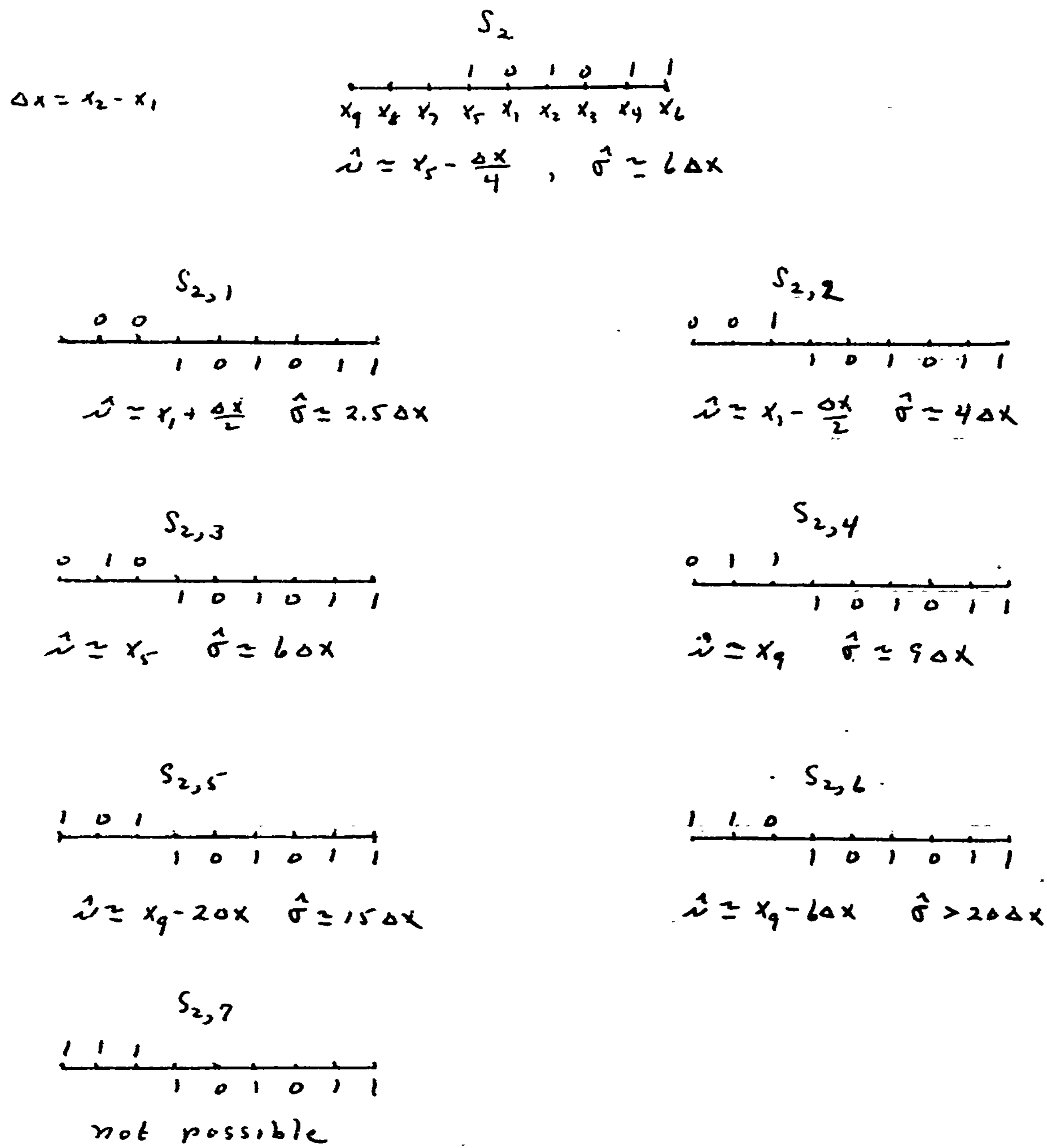


Figure 8, Extension of S_2

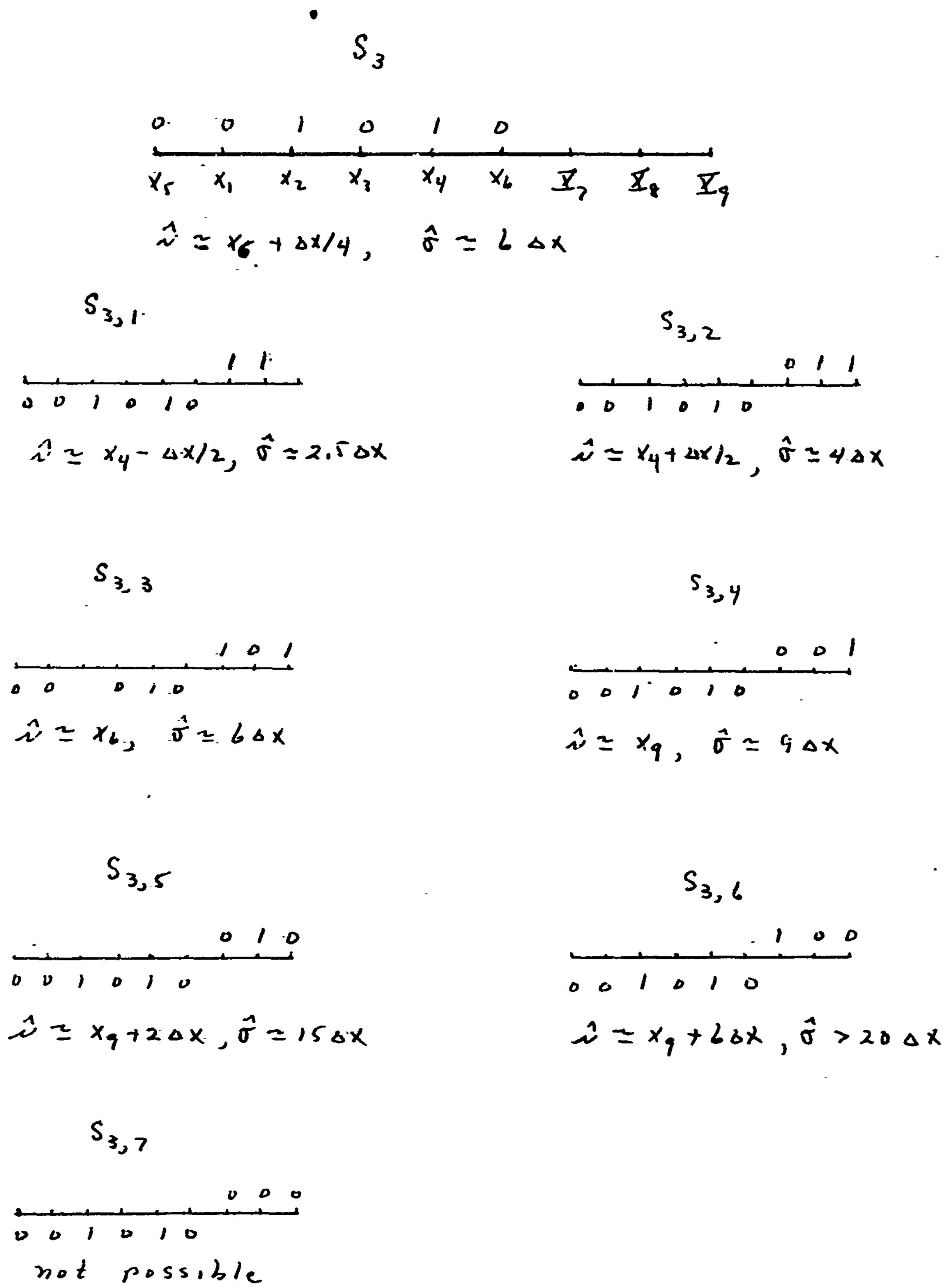
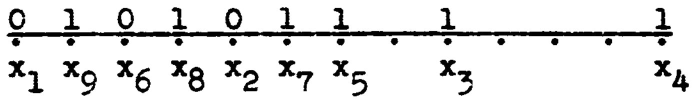
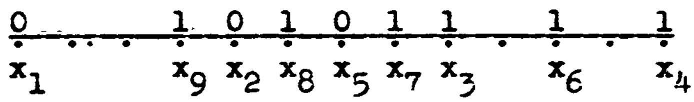


Figure 9. Extension of S_3

The suggested extensions for S_2 and S_3 are motivated by the desire to pin down the ends of the mixed response region by having two '0' responses at the left end of the sample and two '1' responses at the right end. If the responses at the suggested sample points (indicated by x's in Figure 8) do not produce the desired results, additional stimulus levels should be added until this result is achieved. Deviation of the actual results from the desired terminal situation is an indication that σ is still larger than the estimate $\hat{\sigma} \simeq 6\ \Delta x$ for S_2 and S_3. Thus no reduction in step size is warranted. On the other hand, if $\sigma >> 6\ \Delta x$ situations S_2 and S_3 have low probability of occurring and the suggested search procedure will usually terminate with a large Δx. Estimates of μ and σ are included in Figures 7, 8 and 9. Responses at stimulus levels used in the search sequence will have yielded '0' responses at either x_7 or x_9 so that situations $S_{2,5}$ and $S_{2,7}$ will not occur when those responses are considered. This will reduce the number of trials and avoid repetition. Some possible situations where the extension of S_2 could include stimulus levels previously tested are shown in Figure 10. Figure 11 shows a possible boundary effect situations affecting extension of S_2. (The latter is a special case of the situation shown in Figure 10.)

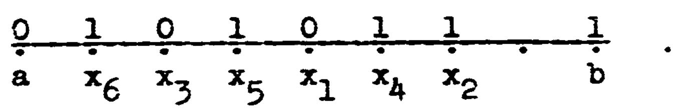


1. Start with $S^* = (0,0,1,1)$ at (x_1,x_2,x_3,x_4)
2. Cycle to S^* at (x_6,x_2,x_5,x_3)
3. Continue to $S_2 = (1,0,1,0,1,1)$ at $(x_9,x_6,x_8,x_2,x_7,x_5)$
4. In Extension of S_2, first point is $x_{10}=x_1$ so $y_{10}=0$



1. Start with S^* at (x_1,x_2,x_3,x_4)
2. Cycle to S^* at (x_2,x_5,x_3,x_6)
3. Continue to S_2 at $(x_9,x_2,x_8,x_5,x_7,x_3)$
4. In Extension of S_2, third point is $x_{12}=x_1$ so $y_{12}=0$

Figure 10. Extensions of S_2 including previous levels



1. Start at (a,b)
2. Cycle to S^* at (x_3,x_1,x_2,b)
3. Continue to S_2 at $(x_6,x_3,x_5,x_1,x_4,x_2)$
4. In Extension of S_2, first point is $x_7=a$ so $y_7=y_8=0$

Figure 11. Extension of S_2 overlapping boundary

VI. Discussion.

The probit method of sensitivity testing requires specification of the stimulus levels to be tested as an input. The staircase technique requires an initial stimulus level and a constant step size in stimulus levels between successive trials as an input. Optimal stimulus level spacing depends upon the distribution of the threshold stimulus and is related to the standard deviation of that random variable. The staircase technique provides some search capability for locating the region of mixed responses but could be inefficient either for the search or for the estimation phases of the investigation. Step sizes which are large in comparison to σ require few steps to locate the mixed response region but have poorer capabilities for estimating μ and σ. On the other hand, step sizes which are small in comparison to σ can require many trials in the search phase if the initial stimulus level is far from the mixed response region. A technique involving reduction in step size is clearly indicated for applications in which the location of the mixed response region is not known a priori and its width may be a small portion of the region to be searched [11].

In specifying the stimulus levels for testing by the probit technique, an approximate location for μ must be known as well as the approximation for σ required for sample spacing. Observations taken outside the region of mixed responses contribute little to the quality of the estimates and may even be hazardous (Berkson [12]). Sample stimulus levels for use with this technique should, ideally, span the region of mixed responses. Large differences between the central stimulus level and μ degrade the estimates (Brownlee et al, [4]).

The assumption that the logarithms (or some other specified transformation) of the threshold stimulus levels are normally distributed provides the basis for the probit technique. The literature contains considerable discussion of this assumption and ways to fulfill or circumvent it (c.f. [1]). For the small sample sizes used in the search phase proposed in this report, the question of normality appears somewhat academic and normality was assumed for the analysis in Sections IV and V. More appropriate distributions may be used for subsequent stages of experimentation.

The factors of safety and reliability involve the tails of the distribution for the threshold stimulus and hence are highly dependent on the mathematical form and parameters of the distribution. For competitive systems wherein it can be assumed that only the parameters of the distribution are different, measures of the location and spread of the distribution such as $\hat{\mu}$ and $\hat{\sigma}$ can be used as the basis for comparison. In estimating absolute levels for safety and reliability, the normal distribution could be used for indicating approximate stimulus levels for subsequent experimentation.

The extensions to the modified binary search suggested in Section V are included so that an investigator can add trials to improve the preliminary estimates of μ and σ given in Section IV if his resources permit and if his resources prohibit a more extensive second stage of experimentation. The results of this extension also provide an indication of the value of the extension. For example, the estimate $\hat{\sigma} = 1.3\ \Delta x$ for S_1 will be replaced by the estimate $\hat{\sigma} = \Delta x$ if the extension of S_1 leads to $S_{1,3}$ and by $\hat{\sigma} = 2.75\ \Delta x$ if the extension leads to $S_{1,4}$. The changes in $\hat{\sigma}$ by extension of S_2 and S_3 are substantially greater thus indicating even further extension is desirable for these situations. The relative frequency of occurrence of S_2 and S_3 should be substantially less than that of S_1.

Subsequent investigations should include a study of the potential variations of $\hat{\mu}$ and $\hat{\sigma}$. Monte Carlo techniques appear more appropriate than an analytic approach for this investigation. The proposed extensions of S_2 and S_3 appear minimal and further extension, at least for these situations, appears desirable.

REFERENCES

[1] Anscombe, F. J., "On Estimating Binomial Response Relations," *Biometrika*, 43 (1956), 461-464.

[2] Bartlett, M. S., "The Use of Transforms," *Biometrics*, 3 (1947), 39-52.

[3] Berkson, J., Review of Reference 7, JASA 43 (1948), 148-153.

[4] Brownlee, K. A., Hodges, J. L., Jr. and Rosenblatt, M., "The Up-and-Down Method with Small Samples," JASA 48 (1953), 262-276.

[5] Dixon, W. J. and Mood, M., "A Method for Obtaining and Analyzing Sensitivity Data," JASA 43 (1948), 109-126.

[6] Evans, F. W. II, MacCarthy, R. C. and Tysver, J. B., A Statistical Design and Analysis for Sensitivity Experiments, Stanford Research Institute Technical Report 001-67, January 1967.

[7] Finney, D. J., *Probit Analysis*, Cambridge University Press, Second Edition, 1952.

[8] Gayle, J. B., "Computer Simulation Study of Bruceton and Probit Methods of Sensitivity Testing," ARO-D Report 64-2, 205-228.

[9] Garwood, R., "The Application of Maximum Likelihood to Dosage-Mortality," *Biometrika* 32 (1941), 46-58.

[10] Golub, A. and Grubb, F. E., "Analysis of Sensitivity Experiments when the Levels of Stimulus Cannot be Controlled," JASA 51 (1956), 257-265.

[11] Langlie, H. J., "A Reliability Test Method for 'One-Shot' Items," Aeronutronics Division of Ford Motor Company, Publication No. U-1792, 1962.

[12] Robbins, H. and Monro, S., "A Stochastic Approximation Method," *Ann. Math. Stat.*, 22 (1951), 400-407.

[13] Tysver, J. B., "Statistical Techniques for Analyzing Data Derived from Experiments on Propellants - I Sensitivity Testing,"Stanford Research Institute Memorandum Report, September 1967.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

Sensitivity Testing for Safety and Reliability

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report - April 1971

5. AUTHOR(S) (First name, middle initial, last name)

J. B. Tysver

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
	34	13

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO.	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

Probit analysis and the "staircase" technique for sensitivity testing require prior knowledge or assumptions to provide the location and/or spacing of stimulus levels to be used in the experimentation. A preliminary experimentation phase using a modification of binary search is proposed to provide the input information necessary for efficient use of these techniques. For applications where the experimentation is severely limited, the results of this modified binary search can be used for the purpose of comparison of the reliability and safety of competitive systems.

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Sensitivity Testing						
Safety						
Reliability						
Binary Search						
Mixed Response Region						
Maximum Likelihood						